FEBS 28025

FEBS Letters 558 (2004) 45–51



# Mutational analysis of GstI protein, a glutamine synthetase translational inhibitor of *Rhizobium leguminosarum*

Chiara Napolitani[a], Luigi Mandrich[b], Anna Riccio[a], Alessandro Lamberti[a], Giuseppe Manco[b], Eduardo J. Patriarca[a,*]

[a] *Istituto di Genetica e Biofisica 'A. Buzzati-Traverso', CNR, Via Marconi 10, 80125 Naples, Italy*
[b] *Istituto di Biochimica delle Proteine, CNR, Via P. Castellino 111, 80131 Naples, Italy*

Received 21 October 2003; revised 11 December 2003; accepted 15 December 2003

First published online 14 January 2004

Edited by Gianni Cesareni

**Abstract** **The small GstI protein (63 amino acids) of *Rhizobium leguminosarum* inhibits the expression of the *glnII* (glutamine synthetase II) gene, thus reducing the bacterial ability to assimilate ammonium. In order to identify the residues essential for its inhibitory activity, all the 53 non-alanine amino acid residues of GstI were individually mutated into alanine. Based on their capacity to inhibit *glnII* expression (in two genetic backgrounds) three groups of mutants were identified. The first group displayed an inhibitory activity similar to the wild-type; the second and the third ones showed partial and total loss of inhibitory activity, respectively. Several mutations of the latter group concerned residues conserved in two related sequences from *Sinorhizobium meliloti* and *Agrobacterium tumefaciens*. Additionally, we performed experiments to exclude a GstI-mediated mechanism of glutamine synthetase II inhibition/degradation. Finally, the protein was over expressed in *Escherichia coli*, purified and characterised.**



*Key words:* Glutamine synthetase; Gene expression; Post-transcriptional inhibition; Alanine-scanning mutagenesis

## 1. Introduction

The ability of *Rhizobium leguminosarum* to assimilate inorganic N depends on two glutamine synthetase (GS) isoenzymes (GSI and GSII) that are differently regulated, at the transcriptional and post-transcriptional level, as a function of the intracellular N status (see [1] for a review). It has been reported that while *glnA* (GSI) transcription is only slightly affected by the N availability, GSI activity is strongly regulated by reversible adenylylation [2,3]. In contrast, it was demonstrated that *glnII* (GSII) transcription, but not GSII activity, is strongly affected by the N status of the cells [4,5]. In particular, it was found that *glnII* transcription is activated (more than 10-fold) under N limitation by the N transcription regulatory system (Ntr system) of signal transduction, comprising the UR/UTase (uridylyl-removing/uridylyltransferase enzyme), the regulatory protein $P_{II}$ and the transcription regulator NtrC [3,6–8]. It has been established that *glnII* expression is also subject to a post-transcriptional mechanism of inhibition that remains largely undefined [9,10]. This mechanism depends on the *gstI* gene that is located upstream and transcribed divergently with respect to *glnII* [9]. *gstI* may act in *cis* or in *trans* with respect to *glnII* not only in *R. leguminosarum* but also in a heterologous genetic background, such as *Klebsiella pneumoniae* reducing their capacity to assimilate $NH_4^+$ [9,10]. Significantly, so far neither GSII nor GstI homologues have been recognised in *K. pneumoniae* as well as in other enteric bacteria. It was observed that when *gstI* is expressed in *R. leguminosarum* incubated under conditions of N excess, low levels of *glnII* mRNA but neither GSII activity nor GSII protein may be detected [4,5,9]. Therefore, GstI was proposed as an inhibitor of *glnII* translation, even though its involvement in a putative mechanism of GSII inhibition/degradation was not excluded. The key role of GstI in *glnII* inhibition was supported by the fact that two out-of-frame mutants (lacking the last 14 amino acids) as well as a single W48D mutant derivative of GstI completely lost their ability to repress *glnII* expression [9,10].

We report here an exhaustive mutational analysis (alanine-scanning) of GstI. The activity of the 53 mutant derivatives obtained was assayed in vivo by measuring their ability to inhibit *glnII* expression in two homologous genetic backgrounds. New perspectives in the study of GstI arise from the availability of the genome sequence (partial or total) of many bacteria. In fact, open reading frames sharing a significant degree of sequence identity with GstI were identified in two phylogenetically correlated bacteria, *Sinorhizobium meliloti* and *Agrobacterium tumefaciens*, thus allowing us to discuss the effect of the mutations in the light of a multisequence alignment. Moreover, in an attempt to define the molecular mechanism of GstI inhibitory activity, we performed experiments whose results support a GstI-mediated mechanism of translation inhibition of *glnII* mRNA and exclude a mechanism of GSII inhibition/degradation.

## 2. Materials and methods

### 2.1. Bacterial strains and media

Strains of *R. leguminosarum* were grown at 30°C on either TYR-rich medium or RMM, a chemically defined medium previously described [10]. *Escherichia coli* strains were grown on TY medium. Antibiotics used were (μg/ml): tetracycline (5), kanamycin (30), rifampicin

*Corresponding author. Fax: (39)-08-17257202.
*E-mail address:* patriarc@iigb.na.cnr.it (E.J. Patriarca).

0014-5793/04/$30.00 
doi:10.1016/S0014-5793(03)01511-4

(100), ampicillin (100). The sources of N used were (mg/ml): glutamine (1) and glutamate (1) or $NH_4Cl$ (1). All media were solidified with 1.5% agar (Difco).

### 2.2. Site-directed mutagenesis

Mutant derivatives of *gstI* were prepared by polymerase chain reaction (PCR) mutagenesis. Plasmid pAR236 (a pGEM7Zf(+) derivative) carrying wild-type *gstI* [9] was used as template for amplification reactions carried out with a high-fidelity *Pfu*-turbo DNA polymerase (Stratagene), using complementary pairs of mutagenic oligonucleotides. PCR reactions were performed as follows: first round 95°C/30 s, followed by 12 cycles of 95°C/30 s, annealing temperature for 1 min, and a final elongation step of 68°C/7 min. An aliquot of the amplified DNA products was digested with *Dpn*I (37°C for 60 min), to degrade the template (methylated DNA). *E. coli* strain DH5α was transformed with the neo-synthesised linear plasmid DNA and from one of the resulting colonies plasmid DNA was purified. The fidelity of the mutagenesis was confirmed by sequencing. DNA sequence was determined with a Thermo Sequenase kit (Amersham) using appropriate synthetic oligonucleotide primers and [α-$^{33}$P]dNTP nucleotides (Amersham).

### 2.3. GstI expression from an inducible promoter

DNAs of plasmids carrying the mutated version of *gstI* were cleaved with *Xba*I and *Nsi*I restriction enzymes, the resulting DNA fragments were separated by electrophoresis, eluted (Qiagen elution kit) and ligated into *Xba*I/*Pst*I-digested pAR231 (a pMP220 derivative) as described elsewhere [10]. By means of this cloning strategy *gstI* as well as its mutant derivatives were cloned under the control of the *dctA* promoter (a succinate-inducible promoter). To exclude the presence of undesired mutations the nucleotide sequence of each mutant was determined. Strains LPR1105 (wild-type) and AR7 (*glnA*$^-$) of *R. leguminosarum* were transformed by conjugation with the help of strain S17-I of *E. coli*. Cells were harvested by centrifugation (3500×*g* for 8 min), washed, mixed and spotted on plates of TYR-agar without added antibiotic. After incubation (16 h, 30°C), the cells were diluted (250 μl TYR) and aliquots were plated on TYR-agar supplemented with rifampicin plus tetracycline (to select LPR1105 derivatives) or with kanamycin plus tetracycline (to select AR7 derivatives).

### 2.4. GS assays and GstI detection by Western blot

To prepare crude extracts *R. leguminosarum* strains were grown in RMM-glucose/glutamine/$NH_4Cl$ (N excess to repress *glnII* transcription) up to exponential growth phase ($OD_{590nm}$ = 0.5–0.6), harvested, washed twice with N-free RMM, and resuspended ($OD_{590nm}$ = 0.3) into RMM-glucose/glutamate (N-limiting to induce *glnII* transcription) supplemented with 20 mM succinate (to induce *gstI* expression). After incubation (3 h, 30°C), bacteria were harvested, resuspended in buffer (20 mM Tris–HCl pH 7.2) and disrupted by sonic oscillation. Cellular debris was removed by centrifugation (12 000×*g*, 4°C, 20 min), and GS activity was measured as described previously [4]. To detect GstI protein Western blots were performed essentially as described elsewhere [10]. Electrophoretic runs were performed with a Mini-Protean II cell unit (Bio-Rad Laboratories) at room temperature. Sodium dodecyl sulphate gel electrophoresis (SDS–PAGE) was performed with gels prepared with a discontinuous gradient (18, 15, and 12.5%) of polyacrylamide. Proteins were transferred onto polyvinylidene difluoride membrane (0.2 mm; Bio-Rad) in buffer (0.375 M Tris-glycine pH 8.3) containing 10% methanol. The membrane was soaked in blocking buffer (phosphate-buffered saline, 5% milk, 0.05% Tween 20), incubated for 1 h with anti-GstI antiserum from rabbit (diluted 1:1000), washed three times for 15 min with blocking buffer and incubated for 1 h with horseradish peroxidase-linked whole antibody (diluted 1:15000) from donkey. The filters were developed by incubation in buffer (0.1 M Tris–HCl, pH 7.6), containing 3,3′-diaminobenzidine (0.3 mg/ml) and $H_2O_2$ (0.01%). Protein concentration was measured using the Bio-Rad protein assay and bovine serum albumin as a standard.

### 2.5. Preparation of total RNA and RT-PCR assays

Total RNA was purified by means of an RNAqueous kit (Ambion). Bacteria (6.5 ml) were harvested by centrifugation (3500×*g*, 4°C, 10 min), resuspended in buffer (10 mM Tris–HCl, 1 mM EDTA) supplemented with lysozyme (1 mg/ml). After incubation (25°C, 5 min) lysis buffer was added and total RNA was purified according to the manufacturer by running the samples onto RNAqueous filter cartridges. Trace DNA contamination was removed using DNA-free DNase (purchased from Ambion). The concentration, integrity and purity of RNA was determined by agarose gel electrophoresis and reading the UV absorbance (260 nm). Reverse transcription (RT) was performed with 0.5 U of MultiScribe Reverse Transcriptase (Applied Biosystems). A *glnII*-specific oligonucleotide 5′-TGCGTCTTGCCACGCAGGTTCGGTACCGGT-3′ (0.4 μM) and total RNA (0.5 μg) were used as primer and template, respectively. An aliquot of the reaction products (3 μl) was used for DNA amplification (PCR) performed using the Amplitaq Gold DNA polymerase (Applied Biosystems) and the same oligonucleotide plus 5′-ATCATCCCTGGCGATCGGATGCGCTTGCG-3′ as primers.

### 2.6. GstI purification, N-terminal sequencing and mass spectrometry analysis

The *gstI* gene of *R. leguminosarum* was cloned in the isopropyl-β-D-thiogalactose (IPTG)-inducible pT7-SCII plasmid by transferring a *Nde*I/*Hin*dIII DNA fragment from plasmid pAR228 [10]. *E. coli* DE3(BL21) cells were transformed with the pT7-SCII-*gstI* construct and cultured in Luria–Bertani medium (100 μg/ml of ampicillin). When the culture reached an $OD_{600nm}$ value of 1, IPTG (0.5 mM) was added. After incubation (4 h, 37°C) the cells (10 g wet weight) were harvested by centrifugation (6000×*g*, 4°C, 10 min), washed with buffer A (25 mM Tris–HCl pH 8.5, 2.5 mM $MgCl_2$, 1 mM dithiothreitol) and stored at −20°C. Finally, cells were thawed and re-dissolved in buffer A (40 ml). Disruption was achieved with a French press cell and after centrifugation (20 000×*g*, 4°C, 20 min) the supernatant was loaded onto a Q Sepharose Fast Flow column (20×2.6 cm; Amersham Pharmacia, Uppsala, Sweden) equilibrated in buffer A and eluted with a linear gradient (0–1 M) of NaCl in buffer A. The fractions containing GstI (identified by Western blotting) were either: (a) pooled and concentrated onto a 10 000 cut-off cellulose membrane (Amicon). The passed through solution was concentrated onto a 3500 cut-off cellulose membrane or (b) concentrated onto a 3500 cut-off cellulose membrane and then loaded onto a a G-75 Superdex column equilibrated and eluted with buffer (25 mM Tris–HCl pH 8.5, 0.5 mM EDTA, 2.5 mM $MgCl_2$, 0.2 M NaCl). The protein analysed by SDS–PAGE appeared, in case a, in the form of a doublet at a molecular mass of about 7000 kDa. An aliquot (∼50 μg) of the purified GstI was made in 0.1% trifluoroacetic acid (TFA) and loaded onto a reverse-phase C18 column equilibrated in 0.1% TFA. Protein was eluted with a gradient of acetonitrile in 0.1% TFA. Four main peaks showing absorbance at both 220 and 280 nm were eluted between 25 and 35% of acetonitrile. After SDS–PAGE analysis, peaks 1, 2 and 4 were analysed by matrix-assisted laser desorption ionisation-time of flight mass spectrometry (MALDI-TOF) and N-terminal sequencing. In case b no doublet was observed but the procedure was performed in the presence of a cocktail of protease inhibitors. Automatic Edman degradation analysis was performed on an ABI Procise protein sequencer, according to the manufacturer's pulse-phase protocol. MALDI-TOF was carried out on a VoayagerDE Pro Biospectrometry Workstation (Applied Biosystems) equipped with a nitrogen laser operated at 337 nm. Protein samples were mixed with a saturated matrix of α-hydroxycinnamic acid and run according to the manufacturer's specifications. CD spectra were recorded with a Jasco J-715 spectropolarimeter equipped with a Peltier-type temperature control system (Model PTC-348WI). The instrument was calibrated with an aqueous solution of D-10-(+)-camphorsulphonic acid at 290 nm [11]. Molar ellipticity per mean residue, $[\theta]$ in deg cm$^2$/dmol, was calculated from the equation: $[\theta] = [\theta]_{obs}$ mrw/10 *IC*, where $[\theta]_{obs}$ is the ellipticity measured in degrees, mrw is the mean residue molecular weight (113 Da), *C* is the protein concentration in g/l, and *I* is the optical path length of the cell in cm. A path length cell of 0.1 cm and a protein concentration of about 0.32 mg/ml were used in the far-UV region. CD spectra were recorded with a time constant of 4 s, a 2 nm bandwidth, and a scan rate of 5 nm/min, were signal-averaged over at least five scans, and baseline-corrected by subtracting a buffer spectrum.

### 2.7. Sequence analysis

Database searches were performed with BLASTP, PSI-BLAST or BLASTPGP (http://www.embl-heidelberg.de). Sequence alignments were done using Clustal X [12].

## 3. Results and discussion

### 3.1. Isolation and activity of GstI mutant derivatives

By PCR reactions performed with overlapping mutagenic oligonucleotides, a collection of 53 GstI derivatives, each one carrying a single amino acid substitution with alanine, were prepared. The activity of the *gstI* mutant derivatives was examined in vivo by measuring the expression of *glnII*, the only recognised target gene, as follows: (i) the mutated genes were cloned under the control of an inducible promoter, *dctA*p, which is induced (more than 20 times) when succinate is added to the growth medium (*dctA* encodes a succinate transporter); (ii) strain LPR1105 (wild-type) of *R. leguminosarum* was transformed with the resulting plasmids including, as a control, plasmid pAR231 (carrying *dctA*p alone) and plasmid pAR237 (carrying the wild-type version of *gstI* under the control of *dctA*p); (iii) the resulting 55 strains were grown in a chemically defined medium (RMM) containing glucose as source of C and glutamine plus $NH_4Cl$ as sources of N (an N excess condition under which *glnII* is transcribed at a very low rate); (iv) the cells were then shifted to RMM-glutamate plus glucose (an N-limiting condition under which *glnII* transcription is 12-fold induced) and with added succinate (to induce *gstI* expression). After 3 h of incubation aliquots of cells were harvested, crude extracts were prepared and GSII activity was measured (see Section 2). In three independent experiments, strain LPR1105-pAR237 (wild-type GstI) showed $175 \pm 25$ U of GSII activity when shifted to RMM with succinate (*gstI* induction), whereas higher values of GSII activity ($700 \pm 100$ U) were obtained with strain LPR1105-pAR231 (vector alone). As shown in Fig. 1, the mutated versions of *gstI* showed a variable behaviour. Based on their capacity to inhibit *glnII* expression (GSII activity is inversely proportional to GstI inhibitory activity) the mutated versions of GstI were arranged in three groups. A first group (group 1), including P2A, G4A, F5A, H6A, R7A, E8A, E27A, R28A, R30A, L31A, E38A, V39A, C40A, D45A, S46A, D58A, and K62A, showed almost the same inhibitory activity (GSII ac-

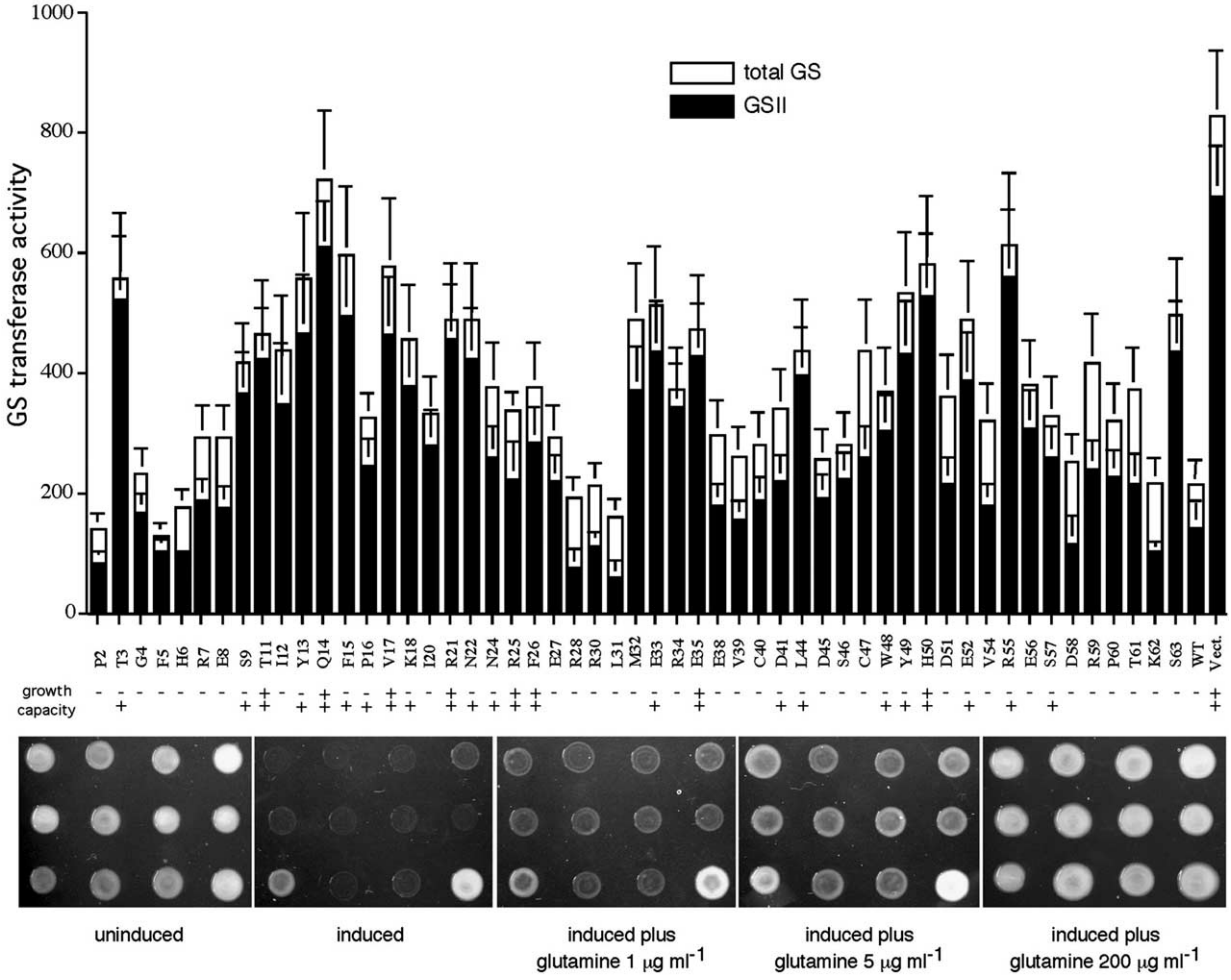


Fig. 1. Effect of GstI mutants on *glnII* (GSII) expression and bacterial growth. Upper part: Transconjugants derived from *R. leguminosarum* (wild-type strain LPR1105) carrying the mutated versions of *gstI* were isolated, grown on RMM-glutamine/$NH_4Cl$ (nitrogen excess, *glnII* repressed) and shifted to RMM-glutamate/succinate (3 h, 30°C), to induce simultaneously *glnII* ($NH_4^+$ deficiency) and *gstI* (succinate). Crude extracts were prepared and assayed for GS activity: total GS activity values (white bars), GSII activity values (black bars), and standard errors are indicated. Lower part: Transconjugants of strain AR7, a *glnA::km* ($GSI^-$) strain of *R. leguminosarum*, were isolated and spotted onto RMM-$KNO_3$ with or without added succinate (*gstI* inducer). Their growth ability was assessed (3 days, 30°C) and is indicated as: (−) undetectable, (+/−) partial, (+/+) normal growth. For instance, the same part of replica plates showing the growth capacity of a group of AR7 derivatives is shown. The same group of mutants spotted on replica plates with RMM-$KNO_3$/succinate but supplemented with glutamine, at the indicated concentrations, is also shown.

tivity lower than 300 U) as compared with the wild-type version of GstI. A second group (group 2) of mutants, including S9A, I12A, P16A, I20A, N24A, R25A, F26A, R34A, D41A, L44A, C47A, W48A, D51A, E52A, V54A, E56A, S57A, R59A, P60A and T61A, was characterised by a partial loss of inhibitory activity (GSII activity between 300 and 450 U). Finally, the last group (group 3) of mutants, including T3A, T11A, Y13A, Q14A, F15A, V17A, K18A, R21A, N22A, M32A, E33A, E35A, Y49A, H50A, R55A and S63A, showed a significantly reduced (GSII activity higher than 450 U) inhibitory activity. The GstI activity was measured under maximal induction of *glnII* transcription (cells shifted from N excess to N deficiency), thus increasing the significance of the results obtained. In fact, usually, GstI exerts its inhibitory activity in *R. leguminosarum* cells growing under conditions of N excess in which GstI is highly expressed whereas *glnII* is transcribed at very low rates [6,9].

The inhibitory activity of GstI mutants was also assessed in strain AR7 of *R. leguminosarum* which harbours a *glnA::km* insertion (GSI$^-$) and thus depends on *glnII* expression (GSII activity) to grow using inorganic sources of N such as nitrate [3]. AR7 derivatives were plated on RMM-nitrate with or without added succinate (*gstI* induction), and their growth ability was compared. As predicted, the mutations that mediate a complete loss of *gstI* activity (as established by measuring GSII activity in the wild-type strain), allow the growth of the *glnA*(GSI$^-$) mutant strain. On the other hand, the addition of glutamine (even at low concentrations) restores a normal growth capacity to those strains that were inhibited after GstI induction, thus excluding a toxic effect induced or provoked by the induction of the mutated versions of GstI (Fig. 1). Some mutated versions of GstI, namely I12A, I20A, R34A and S63A, were able to inhibit the growth of strain AR7(GSI$^-$) although unable to inhibit completely the expression of *glnII* in the wild-type strain. Most likely the ability of strain AR7(GSI$^-$) to induce *glnII* expression is reduced by the absence of GSI. In fact, when spotted on a minimal medium containing $NO_3^-$ as the sole N source, bacteria require GS activity to maintain the intracellular pool of glutamate/glutamine and thus the pool of amino acids (by transamination) needed for gene expression.

### 3.2. *Translation inhibition of* glnII *expression*

To get some insight into the inhibitory mechanism mediated by GstI some strains of *R. leguminosarum* were shifted to a growth condition at which both *gstI* and *glnII* are induced, and used to prepare both total RNA and crude extracts (Fig. 2). RNA samples were subjected to RT-PCR analysis performed with oligonucleotides designed to amplified the 5′-untranslated region (UTR) of *glnII* mRNA whereas crude extracts were used both to detect GstI protein by Western blot analysis, and to measure GSII activity. Specific *glnII* transcripts were detected (Fig. 2A) in all strains analysed, including those expressing: (i) no GstI (carrying the vector alone); (ii) wild-type GstI; (iii) mutated versions of GstI, namely T3A and F15A. With the exception of strain LPR1105-pAR231 (vector alone), the GstI protein was detected in all strains analysed (Fig. 2C). Moreover, the GSII activity measured was variable (Fig. 2B). In fact, as expected, the strain expressing wild-type GstI showed a low level of GSII activity (165 U), while the strains expressing no GstI (vector alone), or the mutated versions T3A or F15A, showed higher levels of GSII

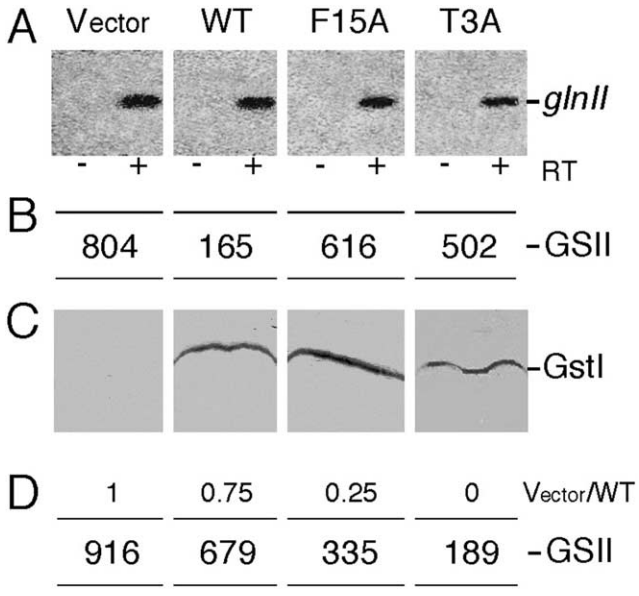


Fig. 2. Post-transcriptional inhibition of *glnII* expression. *R. leguminosarum* strains expressing either wild-type GstI (WT), the mutated versions (T3A or F15A) of GstI, or no GstI (vector alone) were grown under conditions of *glnII* induction and used to prepare total RNA and crude extracts. A: Specific *glnII* transcripts detected by RT-PCR. B: The units of GSII activity measured in the same cells. C: GstI detected by Western blot analysis. D: Crude extracts prepared from *R. leguminosarum* expressing no GstI (Vector alone) or wild-type GstI were mixed at the indicated ratios, incubated for 1 h at 30°C and the GSII activity measured.

activity (500–800 U). Furthermore, specific *glnII* transcripts were detected after induction of other mutated versions of GstI (namely E33A, Y49A, R55A and K62A) that fall both in the central and in the C-terminal region of the protein (data not shown).

A putative role of GstI in a post-translational mechanism of GSII inactivation (covalent modification or protein degradation) was also tested. To this purpose crude extracts obtained from strain LPR1105-pAR237 (expressing wild-type GstI and a low GSII activity) was mixed, at different ratios, with crude extracts prepared from strain LPR1105-pAR231 (undetectable GstI protein and a high GSII activity). After incubation (30°C, 1 h) the GSII activity of the mixtures was measured. As shown in Fig. 2D, the addition of crude extracts containing a high level of the wild-type GstI showed no effect on the GSII activity. In fact, the activity of the mixture decreased but simply as a linear function of the dilution performed. Essentially the same results were obtained by addition of crude extracts containing active or inactive mutated versions of GstI (including T3A, E8A, F15A, E33A, Y49A, R55A and K62A) or by incubating the mixture for a long period of time (up to 3 h) (data not shown). Therefore, we conclude that, at least under the in vitro conditions tested, GstI is unable to induce a post-translational mechanism of GSII inhibition and/or degradation. Instead, our data suggest a block of translation and stabilisation of *glnII* transcripts.

### 3.3. *Identification of GstI homologues*

By using the *R. leguminosarum* GstI sequence (SwissProt accession number Q9K4V1) as a query in a BLASTPGP search in the 'all non-redundant' SwissProt database (release 39.0) we identified putative homologues in two members of


Fig. 3. Multisequence alignment of the *R. leguminosarum* GstI (RlGstI) protein and its putative homologues from *S. meliloti* (SmGstI) and *A. tumefaciens* (AtGstI). Sequences were first aligned with the program Clustal X (gap opening penalty: 10; gap extension penalty: 0; protein weight matrix: identity). Residues identical or conserved in at least three sequences (or two in correspondence of gaps) are shaded and a consensus is reported at the bottom of the figure. Capital letters indicate 100% sequence identity. Color code: basic residues (R, K), red; acidic residues (D, E), violet; polar residues (Q, S, T), green; hydrophobic residues (A, F, V, I, L, W, M), cyan; P, yellow; H and Y, dark green; G, brown; C, pink; N, nothing. Arrows indicate the sites of proteolytic cleavage.

the family Rhizobiaceae, namely *Sinorhizobium meliloti* (GenBank accession number Q8UCR8) and *Agrobacterium tumefaciens* (GenBank accession number Q92TR2) with *E* values of $2\times10^{-14}$ and $7.5\times10^{-2}$, respectively. Moreover, as in the case of *R. leguminosarum*, the putative GstI proteins of both *S. meliloti* and *A. tumefaciens* are encoded by an open reading frame located upstream, and divergent, with respect to a GSII coding gene (data not shown). Thus, hereafter, these proteins will be called GstIs, even though experimental evidence of their involvement in the regulation of *glnII*(GSII) expression is not available.

The sequence alignment (Fig. 3) indicates 46% and 39% sequence identity of RlGstI with SmGstI and AtGstI, respectively, and 35% sequence identity of SmGstI with AtGstI. These values of sequence identity were much lower than those exhibited by orthologous sequences from the same bacteria, including GSII (data not shown). Noteworthy, the GstI proteins exhibit a higher degree of sequence similarity in their N-terminal and C-terminal regions, while they vary more extensively in their central region (Fig. 3). This would suggest that, if the proteins have a similar function, there is a distinct requirement of sequence conservation along their lengths. The invariant residues T3, A10, T11, I12, F15, P16 (N-terminal), A36 (central region), and D45, W48, Y49, H50, A53, K62 (C-terminal) may have a structural and/or functional role. Indeed, it was previously shown that a W48D mutation impaired the inhibitory effect of RlGstI [10]; we demonstrate here that the same holds true for T3, T11, I12, F15, P16, Y49, and H50. In addition, the mutation of residues conserved in two out of the three genes (namely Q14, V17, R21, F26, E33, R34, E35, E52, and R55), also resulted in the loss of GstI activity (Fig. 1). Thus, with few exceptions (such as D45, K62), most of the residues that are conserved among the three genes are required for RlGstI function, whereas only few residues essential for RlGstI activity (such as Y13, N22, and S63) are not conserved. Therefore, on the basis of sequence similarity as well as functional dissection, the possibility that, despite the observed sequence divergences, these proteins may have a similar function cannot be excluded.

Finally, RlGstI and SmGstI gave, in addition, very low similarity scores (BLASTP program) with different DNA and RNA binding proteins. For instance, RlGstI shares similarity with the B subunit of the excinuclease ABC from *Chlamydia pneumoniae* (Domain PD099792 in PFAM; 42% identity/21 residues overlap, *E* value 0.28, covering the N-terminal region of GstI), the single-strand recognition protein of *Drosophila melanogaster* (Domain PD007820 in PFAM; 37% identity/16 residues overlap, *E* value 0.78, covering the N-terminal region of GstI) or the yeast ATP-dependent RNA helicase Dob1 (44% identity/18 residues overlap, *E* value 0.99, covering the C-terminal region of GstI). In correlation, it has previously been proposed that the 5′-UTR of *glnII* mRNA may be the target of the inhibitory mechanism mediated by GstI [9].

### 3.4. GstI purification and characterisation

The GstI protein of *R. leguminosarum* was over-expressed in *E. coli* and purified in two different ways. In a first approach, a combination of ionic exchange chromatography and diafiltration through cellulose membranes (Amicon) was employed. The protein was flowed through a 10,000 cut-off filter while it was concentrated onto a 3500 cut-off filter, thus suggesting a monomeric structure (theoretical MW: 7347). An SDS–PAGE analysis indicated the presence of the pure pro-

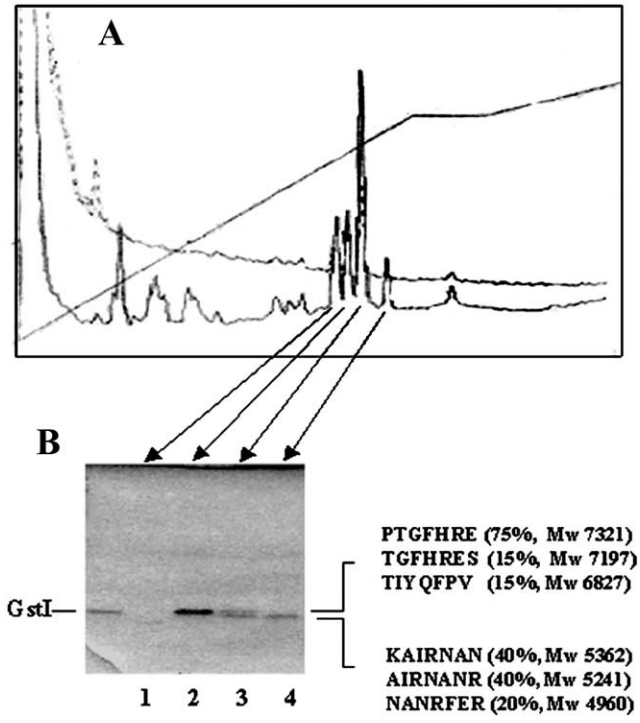


Fig. 4. HPLC analysis of purified GstI. A: HPLC chromatogram of purified RlGstI separated onto a C18 reverse-phase column. Each of the four peaks was analysed by SDS–PAGE (B, lanes 1–4). The masses measured with MALDI-TOF and the N-terminal sequences of peaks 2 and 4 are shown.

tein as a doublet of about 7000 kDa. By using a C18 reverse-phase high performance liquid chromatography (HPLC) column the resulting material was fractionated into four peaks (Fig. 4A) and the analysis by SDS–PAGE (Fig. 4B) indicated three bands of slightly different mass (lanes 1, 2 and 4) in three peaks and one peak that contained two of these bands (lane 3). The three peaks corresponding to lanes 1, 2 and 4 were analysed by mass spectrometry via MALDI-TOF and N-terminal sequencing. The result was no sequence and no mass for peak 1, whereas peaks 2 and 4 were heterogeneous both in mass and in N-terminal sequence. In particular, peak 2 contained the GstI protein with (7321 kDa) or without (7197 kDa) the initial methionine as well as a proteolytic fragment of 6827, corresponding to a truncated version of GstI lacking the first nine amino acids. Peak 4 instead contained masses of 5362, 5241 and 4960 kDa corresponding to GstI without the first 21, 18 and 17 amino acids, respectively. These results indicated partial proteolytic cleavage for RlGstI. Thus a new purification procedure in the presence of protease inhibitors was performed. Essentially, after the anionic step, the protein was concentrated with an Amicon ultrafiltrator (membrane cut-off 3500) and loaded onto a G-75 Superdex column (Fig. 5). The protein was eluted as a symmetrical peak at an elution volume of 88 ml, after cytochrome *c* (MW = 12 500, $V_e = 82.3$) and at a position consistent with a relative mass of about 7400. The protein was identified by SDS–PAGE and Western blot (see inset of Fig. 5). In this chromatographic step a separation occurred between the full-length form (about 7000; peak 6 in Fig. 5) and a form (peak 7) resembling that observed in peak 4 of the HPLC run (Fig. 4). To have an idea about its secondary structure composition the full-length protein was analysed by circular dichroism (Fig. 6). Data analysed with the method of Yang [13] gave the following values: $\alpha = 0$, $\beta = 58.7$, and unstructured = 41.3, allowing us to conclude that GstI has an all-β structure.

In conclusion, a procedure for purification of recombinant

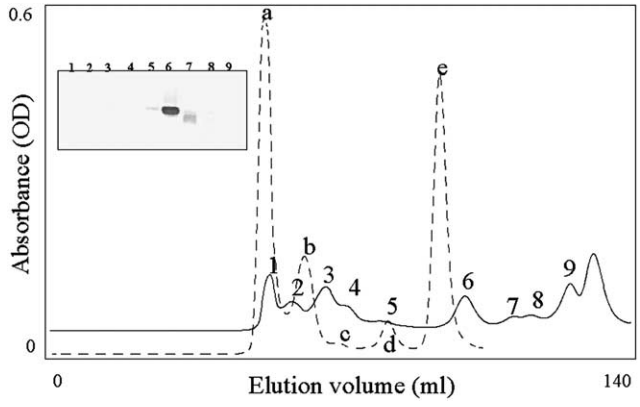


Fig. 5. Gel filtration onto Superdex G-75. The sample from the Q-Sepharose FF, after concentration onto an ultrafiltration apparatus, was loaded and eluted with buffer (25 mM Tris–HCl pH 8.5, 0.5 mM EDTA, 2.5 mM $MgCl_2$ and 0.2 M NaCl), at a flux of 0.5 ml/min (continuous line). A calibration with the indicated molecular weight markers (dashed line) was performed. Markers were as follows: a, blue dextran (>2 000 000); b, bovine serum albumin (66 000); c, ovalbumin (43 000); d, soybean trypsin inhibitor (30 000); e, cytochrome *c* (12 500). In the inset is reported the Western blot (lanes 1–9) of peaks 1–9 indicated in the chromatogram. GstI was eluted in peak 6.

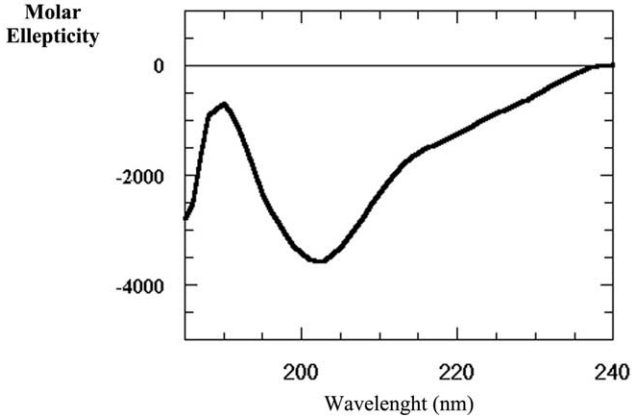


Fig. 6. Circular dichroism of GstI protein. A sample of pure GstI (0.32 mg/ml) diluted in phosphate buffer (40 mM, pH 7) was analysed by circular dichroism in the far-UV region (240–180 nm, 20°C). Data were converted to molar ellipticity.

GstI was devised. The protein appeared particularly sensitive to proteolytic degradation and therefore the presence of protease inhibitors during the purification procedure was mandatory to obtain a full-length protein. Since loops are the protein structures most exposed to the solvent and most prone to proteolytic cleavage, these data suggest that in solution the N-terminus of RlGstI is less structured. The alanine-scanning mutagenesis indicates that this region of the protein plays a functionally important role. In fact, nine out of 16 mutations, significantly affecting GstI activity (namely T3A, T11A, Y13A, Q14A, F15A, V17A, K18A, R21A, and N22A), fall within the N-terminal region of the protein. Thus, even though some mutations could be important per se and not in relation to the structural integrity of the protein, the possibility that the N-terminal region of GstI adopts a functional fold (becomes structured) only in the presence of a ligand (such as an RNA molecule) cannot be excluded [14]. Finally, data obtained by means of SDS–PAGE, MALDI-TOF, gel filtration, and circular dichroism spectrum analysis suggest that GstI is a monomeric protein with an all-β structure. In this context it is important to remark that GstI was proposed to be an RNA binding protein [9] and that the protein domains showing affinity for single-stranded RNA molecules are formed, exclusively, by one or multiple β-sheets [15].

*Acknowledgements:* We thank G. Grimaldi and M. Iaccarino for helpful discussion, L. Camardella and V. Carratore for MALDI-TOF and N-terminal sequence analysis, R. Vito and C. Sole for technical assistance, A. Secondulfo and A. Aliperti for help with the manuscript. The work was supported by MIUR-FIRB (RBNE01KZE7), MURST-CNR Programme L. 488/92 (cluster 02), and Regione Campania. L.M. is a PhD student from the University of Naples 'Federico II'.

## References

[1] Patriarca, E.J., Tatè, R. and Iaccarino, M. (2002) Microbiol. Mol. Biol. Rev. 66, 203–222.
[2] Chiurazzi, M. and Iaccarino, M. (1990) Mol. Microbiol. 4, 1727–1735.
[3] Amar, M., Patriarca, E.J., Manco, G., Bernard, P., Riccio, A., Lamberti, A., Defez, R. and Iaccarino, M. (1994) Mol. Microbiol. 11, 685–693.
[4] Manco, G., Rossi, M., Defez, R., Lamberti, A., Percuoco, G. and Iaccarino, M. (1992) J. Gen. Microbiol. 138, 1453–1460.

[5] Patriarca, E.J., Chiurazzi, M., Manco, G., Riccio, A., Lamberti, A., De Paolis, A., Rossi, M., Defez, R. and Iaccarino, M. (1992) Mol. Gen. Genet. 234, 337–345.
[6] Patriarca, E.J., Riccio, A., Tatè, R., Colonna-Romano, S., Iaccarino, M. and Defez, R. (1993) Mol. Microbiol. 9, 569–577.
[7] Patriarca, E.J., Riccio, A., Colonna-Romano, S., Defez, R. and Iaccarino, M. (1994) FEBS Lett. 354, 89–92.
[8] Schluter, A., Nohlen, M., Kramer, M., Defez, R. and Priefer, U.B. (2000) Microbiology 146, 2987–2996.
[9] Spinosa, M., Riccio, A., Mandrich, L., Manco, G., Lamberti, A., Iaccarino, M., Merrick, M. and Patriarca, E.J. (2000) Mol. Microbiol. 37, 443–452.
[10] Tatè, R., Mandrich, L., Spinosa, M.R., Riccio, A., Lamberti, A., Iaccarino, M. and Patriarca, E.J. (2001) Mol. Plant-Microbe Interact. 14, 823–831.
[11] Venyaminov, S.Y. and Yang, J.T. (1996) in: Circular Dichroism and the Conformational Analysis of Biomolecules (Fasman, G.D., Ed.), pp. 69–107, Plenum Press, New York.
[12] Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) Nucleic Acids Res. 24, 4876–4882.
[13] Yang, J.T. (1986) Methods Enzymol. 130, 208–269.
[14] Barrick, J.E. and Roberts, R.W. (2002) Protein Sci. 11, 2688–2696.
[15] Antson, A. (2000) Curr. Opin. Struct. Biol. 10, 87–94.